How Uniqueness Guides Definite Description Processing

Christopher Ahern and Jon Stevens

1 Introduction

Most analyses of definiteness are based on two important notions: uniqueness and familiarity. Fundamentally, both approaches ascribe some content to the conventional meaning of definite, but not indefinite, descriptions. As an example, consider the following minimal pair.

(1) a. Mary bought the car that she liked.
   b. Mary bought a car that she liked.

Generally speaking, (1a) would only be appropriate if Mary had only liked one car and the hearer was aware of her liking that car. In contrast, (1b) is appropriate if Mary liked more than one car or the hearer was unaware of Mary’s automobile preferences. Thus, the use of the definite, as in (1a), conveys that the referent is both unique and familiar. Much work has gone into testing and refining these theories, the ultimate goal being the simplest theory that provides empirical coverage for the range of uses of the definite. Attempts have been made to make do with only a single notion, deriving the other from pragmatic considerations (cf. Hawkins (1984) in favor of uniqueness and Heim (1982) for familiarity). In this paper we will be interested in bringing evidence from processing to bear on this debate, and testing the possible explanatory role of uniqueness in the processing of definite descriptions.

We explore the effect of determiner choice on listeners’ expectations about possible referents using eye-tracking in the visual world paradigm. The rationale for doing so is the following. First, eye movements to a visual display are closely linked to referential interpretation (Cooper, 1974; Tanenhaus et al., 1995). That is, the location of gaze can be taken as a proxy of what a hearer interprets an expression to refer to. Moreover, eye tracking allows for incredibly precise measurement of the time course of this interpretation. Second, interpretation proceeds incrementally, but is largely predictive. For example, Dahan et al. (2001) find that hearers use frequency to predict the referent of temporarily ambiguous definite descriptions. Participants were presented with a visual scene containing pictures of words containing overlapping initial phonemes (e.g., beetle and beaker) and given instructions to move one of the items on the computer screen. During the temporarily ambiguous instruction, “Pick up the beaker”, hearers were more likely to look to the more frequent of the two items. This and other studies demonstrate the efficacy of eye tracking as a tool to probe expectations during short periods of temporary ambiguity.

The contribution of definites to this predictive process is not entirely clear. For example, Eberhard et al. (1995) find that hearers incrementally use the semantic input from definite descriptions to narrow the set of referential candidates. Participants were presented with visual displays of playing cards arranged on a grid. For different displays, instructions would be disambiguated at different points in time. As a case in point, consider the instructions in (2) with hypothetical points of disambiguation. If there was only one five of hearts below any other card, then the referent would be disambiguated early. If there were two five of hearts below two cards, but only one below an eight, then the referent would be disambiguated in the middle. If there were two five of hearts below two eights, but only one below clubs, then the referent would be disambiguated late.

(2) Put the five of hearts that is in early below the middle eight of late clubs above the three of diamonds

Hearers reliably established the referent only after receiving the disambiguating information. In contrast with the case of lexical ambiguity, there does not seem to be evidence of predictive interpretation. However, unlike the case of lexical ambiguity, hearers have no prior expectations about

*Thanks go to Florian Schwarz for suggestions in design and implementation and to Rachel Stults for running subjects.

which of the two cards would is the more likely referent. It is reasonable to assume that hearers have
very robust information regarding the relative frequencies (e.g., of beetle and beaker), but no prior
expectation that eights are more likely to be clubs, spades, hearts, or diamonds. In fact, it would be
rather surprising if they did.

In what follows, our goal will be to determine whether any sort of prior expectations regarding
the interpretation of definite descriptions can be evoked based solely on the contribution of the
meaning of the definite. In particular, we present hearers with temporarily ambiguous definite de-
scriptions, where a single given referent, which we will call the maximally unique referent, has the
greatest number of unique attributes, i.e. is unique under the greatest number of possible semantic
descriptions. We find that these descriptions, in contrast to their indefinite counterparts, prompt pre-
dictive interpretation in favor of this maximally unique item. The main contribution of this work is a
direct comparison of processing for definite and indefinite descriptions. This comparison allows us
to isolate the contribution of the definite in real time. Additionally, this also brings new evidence to
bear on the theoretical debate; uniqueness is not only a robust notion for describing the meaning of
definites, but also a crucial factor in guiding hearers in the online processing of definite descriptions.

The rest of the paper is organized as follows. In Section 2 we provide a summary of different
approaches to definiteness. On the whole, these considerations suggest that some notion of unique-
ness is essential for an adequate understanding of definites. This requirement leads us to consider
the explanatory role of uniqueness in processing. We proceed to the details of the present study in
Section 3, laying out the predictions made on a uniqueness-based approach. We conclude in Section
4 with a summary of our results and broader connections to theories of meaning.

2 Definiteness

As noted above, most analyses of definiteness involve the important notions of uniquesness and fa-
miliarity. We provide an overview of the two approaches, noting arguments for and against both. Des-
pite the long-standing philosophical interest in definiteness, this discussion will be brief. How-
ever, some notion of uniqueness is unavoidable in a successful theory of definiteness.

2.1 Uniqueness

The uniqueness account of definites stems from Russell’s (1905) influential remarks. Crucially, the
use of a singular definite description is argued to entail the following two conditions. The first
condition requires the existence of the referent; the second requires its uniqueness.

(3) There is an entity in the world satisfying that description.

(4) There is only one such entity.

Under this account, the sentence, ‘The queen of England has hair’ has the following truth conditions.

(5) ‘The queen of England has hair’ is TRUE iff:
   \[ \exists x. queen.England(x) \land (\forall y. queen.England(y) \rightarrow y = x) \land has.hair(x) \]

The sentence is true when there is referent that is unique under the description ‘queen of England’
and that referent has hair.

While this account has been central to subsequent development in the study of definiteness, it is
not entirely unproblematic. Strawson (1950) notes that the existence and uniqueness requirements
cannot be truth-conditional entailments. That is, Russell’s account of the sentence, ‘The king of
France is bald’ makes the following false prediction.

(6) ‘The king of France is bald’ is TRUE iff:
   \[ \exists x. king.France(x) \land (\forall y. king.France(y) \rightarrow y = x) \land is.bald(x) \]

Under this account, the sentence is predicted to be false because there is no king of France. But, as
Strawson notes,¹ we have misgivings about saying that the sentence is false. Rather, it is nonsensical:

¹A similar viewpoint has been argued as implicit in Frege’s analysis of definite descriptions (Beaney, 1997)
we cannot even evaluate the truth or falsity of the sentence. Note that the same would hold if there were multiple kings of France. The sentence would not be false, but again uninterpretable.

In other words, neither the existence nor the uniqueness components of a definite description’s meaning can yield falsity for the sentence containing that definite descriptions. Rather, definites contribute semantic content in the form of presuppositions. The definite article in “the king of France” contributes the presupposition that there is a unique referent in the discourse context that matches the description ‘king of France’.

In a certain sense, definite descriptions seem to be used in ways that accord with both accounts. Donnellan (1966) suggests that definite descriptions can be put to both an attributive and referential use. To take the classic but grisly example, suppose we find Smith horribly murdered. Knowing that he was a sweet and caring person, we might remark that, “The murderer of Smith is insane!” This attributive use conveys the thought that there is some unique individual that murdered Smith, and he or she is insane. Supposing the Jones is arrested and put on trial for the murder, and is behaving erratically at the defense table. We might again say, “The murderer of Smith is insane!” This referential use conveys the thought that Jones is insane, regardless of whether he is the murderer.

Kripke (1977) argues that these two uses of definite descriptions can be reduced to Russell’s account with an appeal to pragmatic principles. Relying on Grice’s (1975) distinction between what is literally said and what is intended to be communicated. Thus, in uttering the proclamation about Smith’s murderer in court we may literally be saying that there is a unique individual who killed Smith and he or she is insane. However, in the context of the courtroom, what we succeed in communicating is the notion that Jones is insane.

Not only does this pragmatic approach resolve the distinction between these apparently different uses of definite descriptions, it also accounts for a broad range of uses noted by Clark (1975). For example, the definite description in (7) does not refer to something that is unique in the world.

(7) I met a man yesterday. The man told me a story.

It is, however, unique when the domain of reference is restricted to the set of entities that are relevant to what’s being said. This restriction can be taken, along the lines of our courtroom example, as an implicature. The pragmatic enhancement of uniqueness-based accounts allows for wide descriptive coverage.

2.2 Familiarity

Familiarity has played a role in analyses of determiners dating from Christophersen (1939). More recently, the influential work of Heim (1982) takes the basic meaning of definites to be the anaphoric uses, as in (7). The domain of reference need not be restricted by external factors, as is the case in uniqueness-based accounts. Heim posits that definites must be used to refer back to a familiar discourse entity, where familiarity is satisfied when an entity has been either explicitly introduced into the discourse (strong familiarity) or implicitly introduced by the context (weak familiarity). Heim (1982) uses the metaphor of file cards, echoing Kamp’s (1981) notion of a Discourse Representation. The use of the different determiners is governed by the following two conditions. The first condition applies to novel referents; the second applies to familiar referents.

(8) For every indefinite description, start a new card.
(9) For every definite, update a suitable old card.

The suitability of old cards might be heavily determined by what is taken to be weakly familiar. After Stalnaker (1974), we can say that referents are at least weakly familiar when their existence is entailed by the common ground of the speaker and the hearer, the common ground being the set of relevant shared-knowledge propositions in a discourse. This notion of weak familiarity explains the following, where the existence of a farmer is entailed by the introduction of a farm into the discourse.

(10) I traveled to the farm, but I couldn’t find the farmer.
The familiarity account also gets at the anaphoric uses of definites, as in the following example of donkey anaphora.

(11) Every farmer who owns a donkey takes the donkey out to dinner.

Here, there is no particular donkey being referred to by the definite DP ‘the donkey’, but rather, for each farmer a new file card is introduced corresponding to the donkey that is owned by that farmer, and the definite description refers anaphorically back to each of those file cards.

Though familiarity is a useful notion for describing anaphora, uniqueness effects are inescapable. Roberts (2003, p.290) notices that sentences like (12) pose a problem for the familiarity account because they refer to discourse-new entities, their only semantic contribution seemingly being the implication of semantic uniqueness.

(12) I opened the door and pushed the button I found inside

Here, the referent of the DP ‘the button I found inside’ is new to the discourse—even weak familiarity does not hold. The implication here is one of uniqueness in the Russellian sense (modulo contextual restrictions on domain). This is a felicitous utterance only when there is a single button inside the box. Roberts unites the uniqueness approach with Heim’s approach by proposing that definites presuppose the existence of a weakly familiar discourse referent that is unique as such, with pure uniqueness effects as in (12) being derived via Gricean implicature: in an out-of-the-blue context it is obviously false that the referent of ‘the button…’ is familiar prior to (12), but the hearer can accommodate a Russellian form of uniqueness in order to make the speaker’s utterance sensible.

If, indeed, some uniqueness component is required to analyze the semantics of definiteness, we may ask whether uniqueness plays a role in the incremental semantic processing of definite descriptions. The results of our experiment suggest that uniqueness does in fact play a role—subjects appear to make predictions about temporarily ambiguous definite DPs, and these predictions tend toward referents which satisfy uniqueness with respect to the greatest number of attributes.

3 Experiment

In this study we examine the effect of uniqueness on the interpretation of definite descriptions. In particular, we want to know whether hearers make rapid probabilistic inferences as an utterance unfolds about the likelihood of possible referents being unique under some as-yet unmentioned description. The following sections show that there is evidence in favor of such predictive inferences.

To clarify the role of uniqueness, consider the temporarily ambiguous definite description ‘the triangle...’ in the context of Figure 1. This description could be disambiguated in various ways. Namely, the disambiguated phrase could be ‘the triangle with the red/green/blue dot,’ or the ‘the triangle with only two equal sides,’ among others. Suppose that shape and color are the only available disambiguating attributes. If color is used as the disambiguating attribute, there are three possible continuations under which there is a unique triangle (‘...with a red/green/blue dot’). Each one of these continuations is equally as likely under a color description. If shape is used, there is only one description that picks out a unique triangle (‘with only two equal sides’). Once a listener has heard ‘the triangle...’, regardless of whether shape or color will be used to disambiguate, the blue triangle is the most probable referent. This is because the blue triangle is not only unique under a color description, but is the only referent that is also unique under a shape description.

We might say that the blue triangle is maximally unique in that it is the felicitous referent of the largest number of definite descriptions. If the definite carries with it the implication of uniqueness, then a listener aiming to maximize the probability of this being satisfied will assume the referent is maximally unique. Thus, if we forced people to guess the intended referent of the ambiguous definite description ‘the triangle’, we would expect most, if not all, to guess the triangle with the blue dot. In contrast, the maximally unique referent is no more likely a referent for an indefinite noun phrase (‘a triangle’) than the others. In fact, we might suppose that an indefinite elicits an expectation that the referent will not be maximally unique. We tested these possibilities by presenting participants with a series of visual arrangements with maximally and non-maximally unique objects and instructions containing both definites and indefinites.
3.1 Materials and Methods

29 undergraduate students at the University of Pennsylvania participated in the experiment for course credit. All participants were native speakers of English. One participant was excluded due to color-blindness. Participants sat at a desk approximately 75cm in front of a computer screen. A chin rest attached to the desk was used to stabilize participants movement and an Eyelink 1000 desk-mounted camera was used to capture gaze direction and duration information. We used the Experiment Builder software package by SR research to present participants with a series of visual arrangements and audio instructions to click on particular objects in the arrangements. For each trial the visual arrangement would appear on the screen, followed by a 500ms pause, and then audio instructions.

There were four target conditions, which corresponded to different sets of instructions. The design was cross-factorial: the two factors were determiner type (definite or indefinite) and whether the referent was maximally unique or not. For example, for the stimulus in Figure 2, the four conditions would be the following:

(13) a. Click on the box that’s next to the triangle with a yellow dot inside.
    b. Click on the box that’s next to the triangle with a red dot inside.
    c. Click on the box that’s next to a triangle with a yellow dot inside.
    d. Click on the box that’s next to a triangle with a red dot inside.

In the first sentence, the definite determiner is used and the referent is maximally unique. In the second, the definite determiner is used, but the referent is not maximally unique. In the third, the indefinite determiner is used and the referent is maximally unique. In the fourth, the indefinite determiner is used and the referent is not maximally unique. Each target stimulus consisted of a maximally unique object, two non-maximally unique objects of the same kind, and a distractor item of a different kind.

We embedded the target noun phrases in a phrasal modifier to ensure that indefinite utterances were as felicitous as definites. Example (14) illustrates the problem: when only one triangle contains a yellow dot, the use of the indefinite article to refer to the triangle yields degraded acceptability.

(14) a. Click on the box that’s next to a triangle with a yellow dot inside.
    b. ?? Click on a triangle with a yellow dot inside.

Each target stimulus consisted of a maximally unique object, two non-maximally unique objects of the same kind, and a distractor item of a different kind. Stimuli were balanced for location, color, and shape. Fillers were constructed so that color and shape we used in equal proportion to disambiguate the referent. This prevented subjects from forming strong expectations regarding what aspects of
Figure 2: Example Stimulus with maximally unique triangle in upper right

the stimuli to attend to. For each target item there was a period of temporary ambiguity between the end of the determiner and a disambiguating post-nominal modifier (‘...the triangle with a...’), during which we tracked the location of subjects’ gaze.

3.2 Predictions

There are two different hypotheses regarding how people behave when processing temporarily ambiguous definite descriptions. We might suppose that people behave in a predictive manner, as if to maximize the probability of a uniqueness being satisfied. We will refer to this possibility as the Predictive Hypothesis. On the other hand, we might suppose that people consider all potentially unique referents equally. We will refer to this as the Equal (Null) Hypothesis. An intuitive notion of what these two hypotheses predict can be seen with regard to Figure 1 in Tables 1 and 2.

If definiteness contributes a meaning of uniqueness and participants predictively interpret definite descriptions, then during the period of ambiguity there will be some increase in the probability of looks to the maximally unique triangle. Let $p$ be some boost that the maximally unique item receives. The probability of looking at the various items over time can be expressed as in Table 1.

In contrast, if definiteness does not contribute a meaning of uniqueness, or if participants do not predictively interpret definite descriptions, then we would expect looks to all three triangles with equal proportion, given that they are all potential candidates for the definite description. The timecourse of these predictions can be seen in Table 2.

Given the design of the task, we should expect a somewhat more subtle behavior. Specifically, participants were asked to identify a particular box, not an individual shape (see Figure 2). This involved scanning not just the objects, but the rows as a whole. Thus, in our analysis we should be concerned with looks to the rows. This leads us to the set of analogous predictions regarding the looks to the row containing the maximally unique element as in Table 3.

In this case we should note that the increase in looks to the maximally unique element and the row containing the maximally unique element could be different. That is, $p$ and $p'$ need not be the same. This follows from the design of the stimuli. Maximally unique elements are twice as likely to appear on the same row as the non-maximally unique objects of the same kind, as compared to the distractor. We controlled for this in the initial design by including an equal number of target stimuli for each condition where the maximally unique object was on the same row as a non-maximally
unique object and the distractor.

3.3 Results

In Figure 3 we plot the results averaged over target trials. Each line represents the advantage of the row with the maximally unique object over the other. This advantage score is equal to the proportion of looks to the row containing the maximally unique object minus the proportion of looks to the other row. Thus, a zero advantage score means both rows are looked to equally. Under the Predictive Hypothesis, we expect that the definite determiner will yield a positive advantage for the maximally unique row. This is borne out—the difference between the determiners is readily apparent, the definite being the only one to favor the row containing the maximally unique element.

We fit a mixed effects model, with subject and item as random intercepts, to determine whether looks to the row containing the maximally unique object depended on determiner choice and time within the period of temporary ambiguity. This period was taken to begin 200ms after the onset of the embedded determiner to allow for the planning of saccadic movement. The period of ambiguity
Table 3: Predictive Hypothesis for Rows

<table>
<thead>
<tr>
<th>REFERENT</th>
<th>...triangle with a...</th>
<th>...yellow</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximally unique row</td>
<td>$\frac{1}{2} + p'$</td>
<td>1</td>
</tr>
<tr>
<td>Non-Maximally unique row</td>
<td>$\frac{1}{2} - p'$</td>
<td>0</td>
</tr>
</tbody>
</table>

Figure 3: Results

was taken to end at the average point of disambiguation. The dependent variable of the model is a binary indicator of whether the subject looked to the maximally unique row during a given 20ms time bin. The results can be seen in Table 4. Crucially, the interaction between time and determiner on looks to the row containing the maximally unique object is significant. We interpret this to mean that the time course of the interpretation of the matrix definite description is sensitive to the definiteness of the embedded determiner. Namely, during the period of ambiguity, the row with the maximally unique object becomes favored over time in the definite condition, but not in the indefinite condition.

3.4 Discussion

The evidence we have presented suggests that the interpretation of definite descriptions is predictive. Given the stimuli, participants had expectations regarding the potential referents of definite descriptions. In particular, their interpretation was guided by uniqueness. When processing definite descriptions, but not indefinites, subjects look reliably more toward the candidate referent that is unique under the greatest number of descriptions. These results show the importance of uniqueness not only in theoretical models, but also in processing behaviors.

Given that there is no increase in the looks to the row containing the maximally unique object in the indefinite condition, we can also rule out any sort ‘pop-out’ effect due to visual saliency. That is, if eye movement revealed an a priori expectation for the maximally unique referent given its visual
Table 4: Mixed Effects Model

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Standard Error</th>
<th>Wald’s z</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>-0.064</td>
<td>0.123</td>
<td>-5.204</td>
<td>1 × 10^{-7}</td>
</tr>
<tr>
<td>Determiner</td>
<td>-0.145</td>
<td>0.077</td>
<td>-1.872</td>
<td>0.0612</td>
</tr>
<tr>
<td>Time</td>
<td>-0.003</td>
<td>0.001</td>
<td>3.899</td>
<td>9.68 × 10^{-5}</td>
</tr>
<tr>
<td>Determiner:Time</td>
<td>0.005</td>
<td>0.001</td>
<td>4.351</td>
<td>1.35 × 10^{-5}</td>
</tr>
</tbody>
</table>

characteristics, then this effect would be seen in both conditions. There is no such effect; both start out with no advantage.

4 General Discussion

We have provided a direct comparison of definite and indefinite descriptions. By allowing a period of temporary ambiguity we were able to isolate the contribution of definiteness to interpretation. Eye movements during this period of ambiguity suggest that the online interpretation of definite descriptions is guided by uniqueness. When processing definites, but not indefinites, subjects look reliably more toward the candidate referent that is unique under the greatest number of descriptions.

These results suggest a role of uniqueness not only in theoretical models, but also in processing behaviors. However, it remains unclear whether uniqueness is exactly the right notion to capture the full range of semantic and psycholinguistic generalizations. Rather, one could posit based on our results that salience, and not uniqueness, plays a special role in interpreting definite descriptions. After all, what we have dubbed the “maximally unique” referent is also the most salient possible referent within its category in that it stands out or “suggests itself”, to echo Schelling (1960). Perhaps interpretation of the definite article triggers an online search for the most salient appropriate referent.

This perspective suggests Lewis (1979, p.348), who asks us to imagine being in a room with a single cat named Bruce, who happens to be causing a stir by running about. We consider the following statement.

The cat is in the carton. The cat will never meet our other cat, because our other cat lives in New Zealand. Our New Zealand cat lives with the Cresswells. And there he’ll stay, because Miriam would be sad if the cat went away.

Initially ‘the cat’ is taken to denote Bruce, but by the end it has come to denote the New Zealand cat. This swap can be taken as a change in the relative salience of the two cats. Uniqueness typically confers a particular kind of salience on a potential referent, and maximal uniqueness a privileged sort of salience. Accounting for uniqueness in these cases as well as salience more generally requires some theory of how our expectations change given ongoing input. The game-theoretic account of definite descriptions presented in Clark (2012) offers a suggestive step towards modeling this more general process. Future research may be brought to bear on whether uniqueness and salience should be distinguished, and on what role these notions play in both psycholinguistic and theoretical analyses of definiteness.

References


